

# Mobile App for Tuberculosis Detection Using Deep Learning and NLP-Based Recommendations

Camily Bravo-Flores , Derik Aranda-Neira , Wilmer Rivas-Asanza , and Bertha Mazon-Olivo 

**Abstract**— Tuberculosis remains one of the leading causes of mortality from infectious diseases and conventional diagnostic methods are often limited by subjectivity. This study presents a mobile application for automated tuberculosis diagnosis, integrating a convolutional neural network (CNN) for chest X-ray classification with a natural language processing (NLP) module for clinical recommendations. Three CNN architectures were evaluated (DenseNet121, ResNet50, and MobileNetV2) using the CRISP-DM methodology, with a Kaggle dataset. DenseNet121 achieved the most balanced performance (Accuracy = 96.2%, Recall = 95.5%, F1-score = 96.2%), prioritizing the reduction of false positives and negatives. The model was integrated into a mobile application developed under the Extreme Programming methodology, employing a client-server architecture with RESTful APIs. The NLP component, based on the DeepSeek-R1 model through the OpenRouter platform, generated contextually relevant clinically responses. Validation with healthcare professionals showed that 91.7% would recommend it as a support tool for tuberculosis diagnosis, due to its diagnostic reliability, usability, and clinical recommendations. These findings confirm the clinical viability and practical utility of the application, positioning it as a valuable resource in healthcare contexts with limited specialist availability.

Link to graphical and video abstracts, and to code: <https://latam.ieceer9.org/index.php/transactions/article/view/10111>

**Index Terms**— Medical diagnosis, artificial intelligence, deep learning, radiography, information technology, tuberculosis.

## I. INTRODUCCIÓN

La tuberculosis es una enfermedad infecciosa que afecta a los pulmones y causa aproximadamente 1.25 millones de muertes anuales según la Organización Mundial de la Salud (OMS) [1]. La mayor cantidad de casos confirmados se concentra en África y Asia; no obstante, en regiones como América Latina, los avances en la detección han sido más limitados, a pesar de la disponibilidad de métodos diagnósticos mejorados [1].

En Ecuador se evidencia una brecha entre los casos de tuberculosis notificados y los estimados por la OMS. Las provincias con mayor incidencia son Guayas (55.03%), El Oro

(7.28%) y Los Ríos (6.02%) [2]. En El Oro se registraron 327 casos en 2021, principalmente concentrados en sectores vulnerables; en Machala, el Hospital Teófilo Dávila reportó una media de 300 casos anuales entre 2019 y 2022, equivalente al 91% de los casos de toda la provincia [3]. La tasa de mortalidad por tuberculosis se encuentra estrechamente asociada a la rapidez del diagnóstico, la cual incide de manera directa en su propagación y transmisión [4].

El presente estudio realiza una revisión técnica de los modelos de aprendizaje profundo aplicados al análisis de imágenes médicas, así como de los sistemas de recomendación clínica basados en inteligencia artificial. El proceso metodológico comprende desde la definición y comprensión del problema y la preparación de los datos, hasta el entrenamiento, comparación y evaluación de distintos modelos, lo que finalmente conduce a la selección de DenseNet121 como modelo implementado para la detección de tuberculosis.

### A. Estado del Arte

#### 1) Deep Learning

Investigaciones recientes han evidenciado el potencial de modelos de aprendizaje profundo y de procesamiento del lenguaje natural (NLP). Rahman et al. [5] reportan arquitecturas de redes neuronales convolucionales (CNN) como ResNet, VGG, Inception y DenseNet, combinadas con técnicas de segmentación pulmonar como U-Net, alcanzan valores AUC cercanos a 1.0 en la detección de tuberculosis. Por otro lado, Kotei y Thirunavukarasu [6] propusieron un modelo híbrido con Vision Transformers y ResNet, capaz de capturar relaciones espaciales de largo alcance y mejorar la precisión. Además, varias técnicas de aprendizaje profundo [7], en las que arquitecturas como ResNet y variantes de redes preentrenadas como EfficientNet predominan, constituyen la mayoría de los enfoques reportados. Prieto-Ordaz et al. [8] integraron CNN (ResNet, DenseNet) con LSTM para generar reportes médicos a partir de radiografías, automatizando tareas tradicionalmente dependientes de radiólogos. Finalmente, Guanca et al. [9] evaluaron enfoques tradicionales como SVM (C=10), Random Forest (1000 árboles) y CNN simples de tres capas, obteniendo también bajos índices de error.

#### 2) Procesamiento de Lenguaje Natural (NLP)

Pupo, Nuñez y Pino [10] evidencian la utilidad de NLP para extraer información clínica relevante, automatizar revisiones de literatura, identificar patrones en historiales médicos y generar hipótesis diagnósticas, contribuyendo así a la reducción de los tiempos de análisis. Por otra parte, Guerrero et al. [11] enfatizan la necesidad de incorporar principios éticos y procesos de validación rigurosos para garantizar interacciones confiables entre los sistemas automatizados y los

The associate editor coordinating the review of this manuscript and approving it for publication was Samuel Ortega (*Corresponding author: Bertha Mazon-Olivo*).

C. Bravo-Flores, and D. Aranda-Neira, are with Facultad de Ingeniería Civil, Universidad Técnica de Machala, Machala, Ecuador (e-mails: cbravo8@utmachala.edu.ec, and daranda2@utmachala.edu.ec).

W. Rivas-Asanza, and B. Mazon-Olivo are with Grupo de investigación AutoMathTIC, Facultad de Ingeniería Civil, Universidad Técnica de Machala, Machala, Ecuador (e-mails: wrivas@utmachala.edu.ec and bmazon@utmachala.edu.ec).

profesionales de la salud.

Estudios en otros dominios refuerzan este potencial: Jiménez et al. [12] aplicaron modelos preentrenados (EfficientNetB0, ResNet50, VGG19) para detectar enfermedades foliares en banano, ampliando un dataset de 900 a 9,000 imágenes mediante *data augmentation* y mejorando la toma de decisiones agrícolas. Por otro lado, Guncay et al [13] alcanzaron una mayor precisión y capacidad de generalización en el diagnóstico de enfermedades bucales mediante la optimización de hiperparámetros y el aumento de datos, lo que evidencia la aplicabilidad clínica de la inteligencia artificial.

En conjunto, la literatura evidencia que el aprendizaje profundo y el NLP constituyen tecnologías clave para la optimización diagnóstica y la toma de decisiones clínicas.

### B. Contribuciones

Actualmente existen avances tanto en el diagnóstico automatizado de enfermedades, como en el uso de asistentes virtuales basados en NLP. Sin embargo, no existe una solución que integre ambos enfoques en una única herramienta especializada. En este contexto, las principales contribuciones de este trabajo son las siguientes:

1. **Diseño de una Arquitectura Híbrida para Diagnóstico y Asistencia.** Se propone la integración de modelos CNN y técnicas de NLP en una aplicación móvil capaz de diagnosticar tuberculosis a partir de radiografías de tórax y ofrecer asistencia conversacional en tiempo real, adaptada a cada caso ingresado por el usuario.
2. **Evaluación Comparativa de Modelos CNN.** Se realiza un análisis comparativo de los modelos DenseNet121, ResNet50 y MobileNetV2, utilizando métricas de rendimiento para identificar la arquitectura más estable y precisa para la detección de tuberculosis, considerando tanto su desempeño como la facilidad de implementación en entornos con recursos computacionales limitados.
3. **Desarrollo de una Aplicación Accesible.** Se presenta el diseño de una aplicación móvil ligera, orientada a operar en contextos con restricciones de infraestructura tecnológica. La herramienta permite la carga de radiografías de tórax, realiza el análisis de probabilidad de tuberculosis mediante DenseNet121 y genera una recomendación médica. Adicionalmente, incorpora un chatbot basado en NLP, entrenado con información médica sobre tuberculosis, que permite formular preguntas y recibir orientaciones personalizadas. Este flujo automatizado contribuye a reducir la brecha diagnóstica en regiones vulnerables y ofrece una experiencia accesible incluso en ausencia de personal médico especializado, constituyéndose en una herramienta de apoyo para profesionales de la salud.

En síntesis, se realizó una evaluación comparativa de las arquitecturas DenseNet121, ResNet50 y MobileNetV2, identificando la configuración más estable y eficiente en términos de precisión diagnóstica y costo computacional. Este resultado permitió integrar un modelo de deep learning y NLP en un entorno móvil accesible, configurando una arquitectura híbrida que apoya el diagnóstico automatizado de

enfermedades infecciosas. La propuesta representa un avance tecnológico para sectores rurales con acceso limitado a especialistas y equipamiento médico.

## II. REVISIÓN DE LITERATURA

**Deep Learning.** Es una rama del aprendizaje automático que emplea redes neuronales artificiales en múltiples capas para modelar y extraer representaciones jerárquicas de datos. Su capacidad para identificar patrones complejos lo convierte en una técnica adecuada para el análisis de datos no estructurados, tales como: imágenes, audio y texto. En particular, ha demostrado un alto rendimiento en el contexto de análisis de imágenes médicas contribuyendo a un diagnóstico más preciso [14]. La técnica de *transfer learning* permite aprovechar el conocimiento adquirido por modelos previamente entrenados con grandes conjuntos de datos, como *ImageNet* y transferirlo a nuevas tareas como la clasificación de imágenes médicas. Basado en este enfoque, el método de *Feature Extractor*, congela la parte convolucional del modelo y entrena únicamente el clasificador final, mientras que *Fine-tuning*, aplica un ajuste en los pesos del modelo para una adaptación más precisa al nuevo dominio de datos [15]. Jacobo-Zavaleta y Zavaleta [16], destacan que estas técnicas superan limitaciones del diagnóstico tradicional.

Por lo tanto, la selección de las arquitecturas se justifica por su desempeño en extracción jerárquica de características y modelar patrones complejos en entornos médicos. Por ejemplo, ResNet incorpora conexiones residuales que facilitan el entrenamiento mitigando problemas de desvanecimiento con estabilidad y capacidad de generalización. Asimismo, DenseNet maximiza la reutilización de información mediante conexiones directas entre capas. Finalmente, MobileNet emplea convoluciones separables en profundidad, lo que permite un balance óptimo entre rendimiento y eficiencia computacional. En conjunto, estas arquitecturas ofrecen presión diagnóstica, eficiencia computacional y escalabilidad en sistemas de apoyo clínico en plataformas móviles orientadas al diagnóstico clínico.

**Procesamiento de Lenguaje Natural (NLP).** Es una rama de la inteligencia artificial que permite a los sistemas computacionales comprender, interpretar y generar lenguaje humano facilitando tareas como la extracción de información y la generación de texto en contextos específicos como el clínico [17]. Los sistemas NLP se emplea en sistemas de recomendación médica para sugerir tratamientos y seguimiento, procesando grandes volúmenes de información [18]. Por su parte, Terraza et al. [19] destacan la capacidad de NLP para interpretar señales complejas y modelar información secuencial, permitiendo a los sistemas comprender, estructurar y generar contenido clínico útil.

En este proceso se tomó en cuenta la disponibilidad de frameworks ligeros combinados con plataformas en la nube facilitan la integración entre el análisis de imágenes médicas e interfaces de usuario. Los sistemas basados en arquitecturas RAG permiten conectar los entornos de producción con modelos avanzados de lenguaje para generar respuestas contextualizadas que se adaptan al usuario [20].

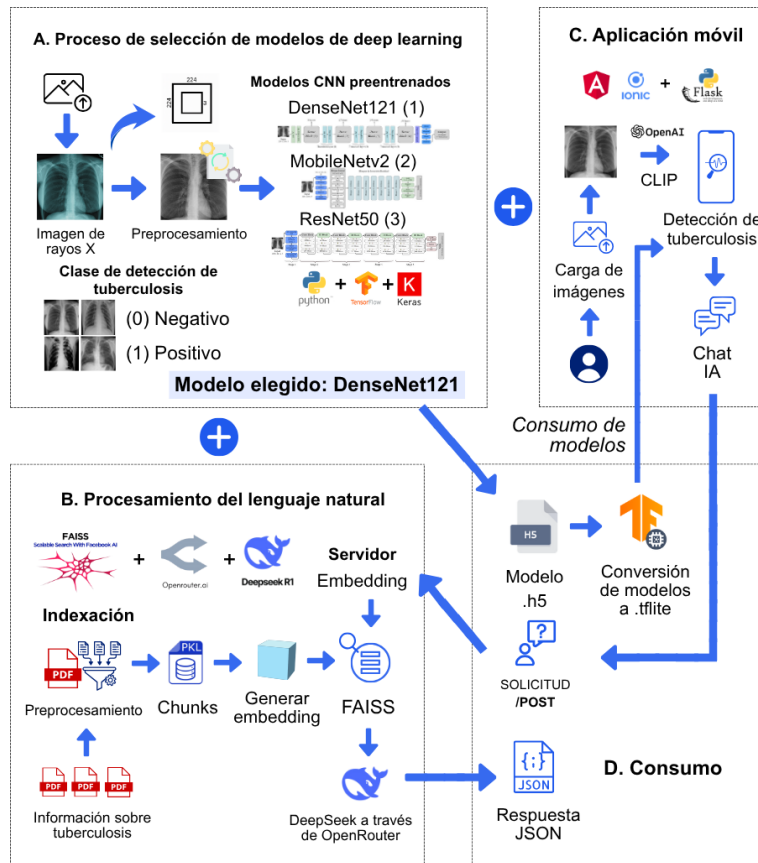


Fig. 1. Arquitectura de procesos para el entrenamiento de modelos CNN y procesamiento de NLP para su consumo.

### III. PROPUESTA DE ARQUITECTURA DEL SISTEMA

La Fig. 1 presenta la arquitectura del sistema desarrollado. Integra cuatro componentes: A) Selección del modelo de deep learning para la detección de tuberculosis en radiografías de tórax, B) Procesamiento de lenguaje natural para generar recomendaciones médicas sobre la enfermedad, C) Aplicación móvil que diagnostica tuberculosis en imágenes de tórax y asiste al usuario mediante un chatbot incorporado, y D) Consumo del modelo CNN elegido mediante API REST. En el desarrollo de los componentes A) y B) se aplicó la metodología CRISP-DM, con sus fases: comprensión del negocio, preparación de los datos, modelado, evaluación y despliegue [21], [22]. En C), desarrollo de la aplicación móvil, se utilizó la metodología Extreme Programming (XP), por ser ágil y útil para integrar funcionalidades de forma continua e iterativa, desde la carga de imágenes hasta la presentación de resultados y recomendaciones [23]. A continuación, se describen cada componente.

#### A. Proceso de Selección de Modelos de Deep Learning

Este proceso (Fig. 1.A) fue realizado según la metodología CRISP-DM; sus fases son:

**Comprensión del Negocio.** En el ámbito médico, la inteligencia artificial y el Deep Learning son tecnologías clave para diagnósticos precisos y eficientes. Se estudió modelos de clasificación CNN enfocados en la detección temprana de enfermedades, con énfasis en tuberculosis a partir de imágenes

radiográficas.

**Comprensión de los Datos.** Como se observa en la Fig. 2, se recopilaron imágenes de radiografías de tórax, con etiquetas de diagnóstico de tuberculosis, de diferentes conjuntos públicos en Kaggle [5], [24]. Estos conjuntos fueron integrados y para garantizar un equilibrio entre las clases de tuberculosis (Negativo y Positivo), se conformó un dataset de 5,800 imágenes, descartando imágenes sobrantes de manera controlada. En la Fig. 2, se muestran ejemplos de imágenes de ambas clases.

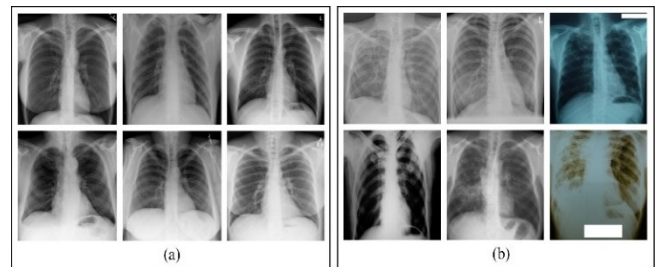


Fig. 2. Radiografías con casos de tuberculosis Negativo (a) y Positivo (b) de [24].

**Preparación de los Datos.** El preprocesamiento incluyó el redimensionamiento a  $224 \times 224$  píxeles, recorte en la región pulmonar y normalización en un rango de  $[0, 1]$ . Adicionalmente, se aplicaron técnicas de *data augmentation* tales como la rotación, el desplazamiento, zoom, ajuste de

brillo, volteo horizontal e inclinación, aumentando el dataset a 11,600 imágenes. Con el objetivo de simular la variabilidad de uso en entornos reales, incrementando así la robustez y generalización del modelo para evitar el sobreajuste como se indica en la Tabla I.

TABLA I  
TÉCNICAS DE AUMENTO DE DATOS APLICADAS

Transformación	Aplicación
Rotación	$\pm 20^\circ$
Desplazamiento	Horizontal y vertical ( $\pm 20\%$ )
Zoom	$\pm 20\%$
Ajuste de Brillo	0.8 a 1.2
Volteo Horizontal	Aplicado de forma aleatoria
Inclinación	$\pm 15^\circ$

En la Tabla II, se detalla la división del conjunto de datos que contiene 11,600 imágenes para el estudio, en tres subconjuntos principales: entrenamiento (80%), validación (15%) y prueba (5%).

El preprocesamiento y la ampliación de datos se llevaron a cabo utilizando marcos como TensorFlow y Keras [25], que ayudan a la aplicación de técnicas avanzadas de optimización para la generalización en modelos de aprendizaje profundo en imágenes médicas [26]. El entrenamiento se efectuó en un laptop con Windows 11, implementada con recursos hardware especificados en la Tabla III.

TABLA II  
DISTRIBUCIÓN DEL CONJUNTO DE DATOS

Clases de tuberculosis	Entrenamiento 80%	Validación 15%	Prueba 5%
Negativo	4,640	870	290
Positivo	4,640	870	290
Total	9,280	1,740	580

TABLA III  
CONFIGURACIÓN DEL AMBIENTE EXPERIMENTAL

Sistema Operativo	Windows 11 Home, v. 24H2
Procesador	Intel® Core™ i9-14900HX @ 2.20 GHz
Memoria RAM	32 GB @ 5,600 MHz
Tarjeta gráfica	NVIDIA GeForce RTX 4060 Laptop GPU (8 GB VRAM)
Almacenamiento	1 TB

Los modelos CNN fueron evaluados mediante métricas con sus valores clave: verdaderos positivos (TP), radiografías con tuberculosis correctamente detectadas; verdaderos negativos (TN), radiografías sanas clasificadas correctamente; falsos positivos (FP), casos sanos clasificados erróneamente como positivos; y falsos negativos (FN), casos con tuberculosis no detectados. Con base en estos valores, se calcularon las siguientes métricas de evaluación:

- *Accuracy* es la proporción de predicciones correctas en relación con el total de predicciones realizadas [27].

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \quad (1)$$

- *Recall* es conocido como *sensitivity*, es la capacidad del

modelo para detectar correctamente los casos positivos [27].

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

- *Specificity* es la capacidad de identificar correctamente los casos negativos [27].

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

- *Youden's index* permite evaluar la capacidad global de discriminación de un modelo de clasificación binaria para determinar la distinción entre enfermos y no enfermos [27].

$$J = Recall + Specificity - 1 \quad (4)$$

- *F1-Score* conocido como la media armónica entre precisión y sensibilidad; equilibra ambos valores [28].

$$F1 - Score = 2 \times \frac{Accuracy \times Recall}{Accuracy + Recall} \quad (5)$$

- *Precision* es la proporción de verdaderos positivos entre todas las predicciones positivas [29].

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

- Área Bajo la Curva ROC (*AUC-ROC*), mide la capacidad del modelo para distinguir entre clases [28].

En este estudio, se determinó que la métrica más relevante es el *Recall*, por su capacidad para identificar todos los casos positivos de tuberculosis y minimizar falsos negativos, reduciendo riesgos clínicos graves.

**Modelado.** Se evaluaron tres arquitecturas de CNN preentrenadas, seleccionadas según características acordes a los objetivos del proyecto, que se describen a continuación.

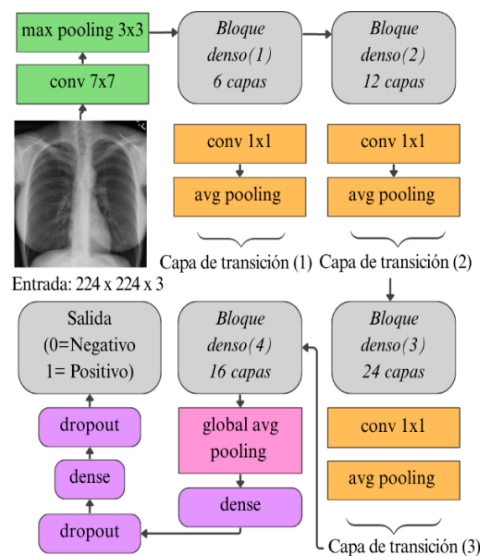


Fig. 3. Arquitectura DenseNet121 utilizada para la detección de tuberculosis en radiografías de tórax.

*DenseNet121* (Fig. 3) está compuesta por cuatro bloques densos (6, 12, 24 y 16 capas) donde cada capa recibe como entrada las salidas anteriores del mismo bloque. Incluye capas de transición, *global average pooling* y una capa densa con *dropout* (40%). La salida utiliza una función de activación *sigmoide* para clasificar la enfermedad de tuberculosis

Negativo y Positivo. Esta arquitectura ha mostrado alto rendimiento en el diagnóstico médico, alcanzando un *Accuracy* del 91%, *Recall* del 90% y *F1-score* del 89% [30].

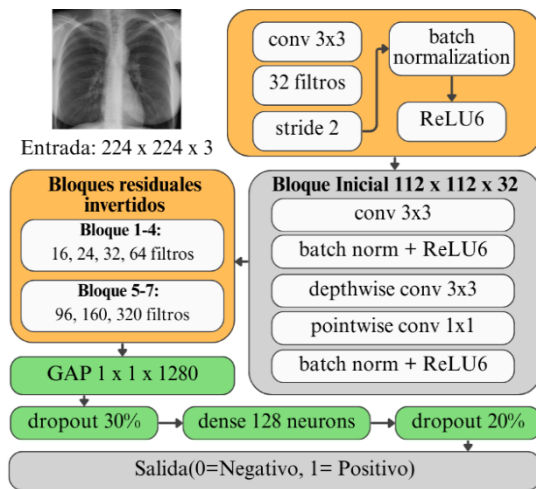


Fig. 4. Arquitectura MobileNetV2 utilizada para la detección de tuberculosis en radiografías de tórax.

*MobileNetV2* (Fig. 4.) es una arquitectura optimizada para dispositivos móviles y entornos con recursos limitados. Basada en bloques convolucionales invertidos con conexiones residuales lineales que permiten un desempeño competitivo [31]. Se inicia con una capa de entrada de imágenes ( $224 \times 224 \times 3$ ), seguida de una capa de convolución  $3 \times 3$  con *stride 2* para reducir dimensiones, seguida de *Batch Normalization* y una función de activación *ReLU6* para optimizar la extracción de características. Luego, se organiza en bloques de Inversión Residual con convoluciones *depthwise* y *pointwise*. Finaliza con *global average pooling*, una capa densa de 128 neuronas, *dropout* para regularización y salida *sigmoide* binaria (0 = Negativo, 1 = Postivo).

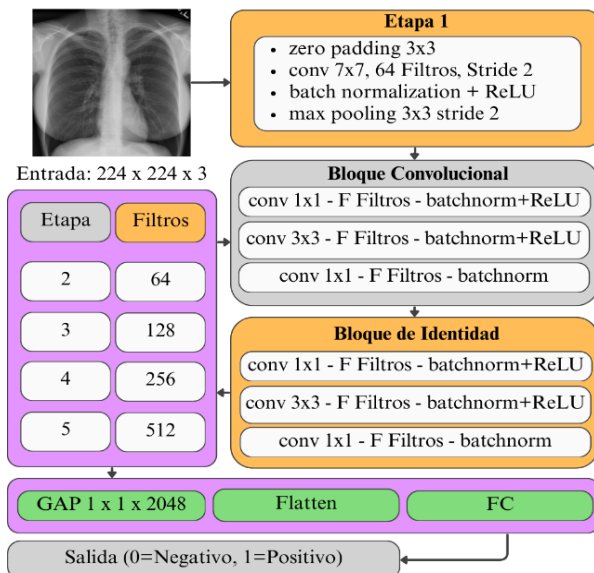


Fig. 5. Arquitectura ResNet50 utilizada para la detección de tuberculosis en radiografías de tórax.

*ResNet50* (Fig. 5), se caracteriza por el uso de conexiones residuales que capturan características jerárquicas de las

imágenes. Esta estructura es capaz de extraer características altamente discriminativas, logrando un desempeño notable en el diagnóstico de enfermedades pulmonares, incluida la tuberculosis, lo que ha sido validado en estudios recientes que destacan su precisión y robustez en aplicaciones clínicas [32]. Se utilizó una parametrización preestablecida, para lograr resultados estandarizados en todas sus pruebas.

TABLA IV  
CONFIGURACIÓN DE HIPERPARÁMETROS

Parámetro	Valor
Input size	$224 \times 224$ píxeles
Batch size	32
Number of epochs	40
Optimizer	Adam
Learning rate	0.0001
Include top	False
Weights	ImageNet
Pooling	Global Average Pooling
Frozen layers:	80% en todas las arquitecturas
Final activation	Sigmoide (clasificación binaria)

En la Tabla IV, se presentan los hiperparámetros aplicados a los tres modelos CNN con los que se realizó la comparación en este artículo.

### B. Procesamiento de Lenguaje Natural

Este componente de NLP (Fig. 1.B) se orientó en generar recomendaciones clínicas contextualizadas a partir de los resultados del diagnóstico. Se aplicó CRISP-DM, para una implementación sistemática y escalable del asistente conversacional mediante fases de desarrollo, las cuales se describen a continuación:

**Comprensión del Negocio.** Abordó el asesoramiento en la interpretación de resultados clínicos mediante un sistema NLP, que entrega recomendaciones automatizadas y contextualizadas. En suma, facilitando la comprensión del diagnóstico y aportando valor a la atención médica digital.

**Comprensión de los Datos.** Se recopilaron documentos técnicos que conforman la base de conocimiento, una síntesis de fuentes científicas y dos referencias oficiales: el manual técnico de la OMS sobre tuberculosis [33] y el *Core Curriculum on Tuberculosis* de los Centros para el Control y la Prevención de Enfermedades (CDC) [34]. Estos documentos contienen información validada sobre síntomas, diagnóstico, tratamiento y prevención, permitiendo al sistema generar respuestas clínicas fundamentadas y coherentes.

**Preparación de los Datos.** Los textos fueron procesados y normalizados, eliminando ruido y unificando formatos. Se utilizó modelos preentrenados de *Sentence Transformers* para convertirlos en vectores semánticos mediante *embeddings* basados en arquitecturas *transformer*. Con estos vectores se construyó un índice de recuperación eficiente con FAISS para localizar rápidamente fragmentos relevantes dentro de grandes volúmenes de datos para garantizar respuestas precisas y oportunas [35].

**Modelado.** Se integró el modelo de lenguaje de código abierto DeepSeek R1 (arquitectura *Mixture of Experts*, que activa solo una fracción de sus parámetros por tarea,

optimizando recursos sin perder rendimiento) y el índice FAISS como núcleo de un sistema Retrieval-Augmented Generation (RAG) [36]. RAG incorpora conocimiento externo, recuperación de información y generación de respuestas precisas, actualizadas y adaptadas al dominio del proyecto. Por tanto, superando las limitaciones de modelos entrenados con datos estáticos. DeepSeek R1 receipta la petición del usuario, analiza los documentos recuperados y produce respuestas en lenguaje natural adaptadas al dominio clínico. En pruebas piloto, se evaluaron ejemplos reales, ajustando hiperparámetros como número de documentos recuperados, longitud máxima de respuestas y umbrales de similitud; logrando un equilibrio entre precisión y eficiencia. El resultado es un modelo híbrido capaz de comprender consultas y ofrecer respuestas concretas, relevantes y contextualizadas en el dominio clínico, fortaleciendo la calidad informativa del asistente conversacional.

### C. Desarrollo de Aplicativo Móvil

El tercer componente corresponde al desarrollo de la aplicación móvil (Fig. 1.C) utilizando el framework Ionic, por su capacidad multiplataforma, lo que permite a los usuarios interactuar con el sistema de diagnóstico de tuberculosis y recibir recomendaciones médicas. Para su implementación se aplicó la metodología ágil Extreme Programming (XP), lo que facilitó la integración progresiva de funcionalidades clave, asegurando una experiencia fluida, validación continua y rápida adaptación a los cambios. El desarrollo se organizó en las siguientes iteraciones:

- **Iteración 1:** Implementar un proceso de carga y validación de imágenes radiográficas de tórax y su descripción textual, mediante el modelo CLIP (Contrastive Language-Image Pre-training) de OpenAI, con el objetivo de evaluar imágenes ingresadas por el usuario y aceptar como válidas aquellas con una similitud mayor o igual al 80% al tipo de imagen objetivo [37].
- **Iteración 2:** Conexión de la aplicación móvil con el modelo CNN en formato TensorFlow Lite (TFLite), seleccionado por sus mejores métricas, integrando las metodologías CRISP-DM y XP.
- **Iteración 3:** Incorporación de la funcionalidad de diagnóstico, que analiza una imagen y devuelve una probabilidad de tuberculosis. Además, se realizó validación mediante revisiones de código y pruebas de aceptación.
- **Iteración 4:** Pruebas de rendimiento de la aplicación, evaluando tiempos de respuesta, fluidez y accesibilidad, con mejoras continuas. Se realizó la integración de chatbot basado en NLP.
- **Iteración 5:** Documentación del proyecto guiada por los principios de XP, priorizando simplicidad, comunicación constante y adaptabilidad.

Como parte del proceso de evaluación para este componente, se establecen criterios orientados a medir la calidad de interacción y el grado de utilidad de la información ofrecida mediante el uso de encuestas al personal médico. Considerando usabilidad, facilidad de interpretación, pertinencia y utilidad de las recomendaciones generadas por el sistema conversacional.

### D. Consumo

Este componente (Fig. 1.D), facilita la interacción e intercambio de información entre la aplicación móvil y el servicio de inteligencia artificial (diagnóstico y recomendación), mediante una Interfaz de Programación de Aplicaciones (API), de tipo *RESTful* (REpresentational State Transfer) [38].

El modelo de diagnóstico se implementó como un servicio web backend desarrollado en Python e integrado en una arquitectura web basada en la nube para asegurar disponibilidad y escalabilidad. Este enfoque es consistente a otras plataformas de análisis de imágenes médicas y con implementaciones de Flask en entornos clínicos [39], [40]. La aplicación móvil envía las radiografías a través de la API REST al modelo, que se encarga de realizar la inferencia y retornar la clasificación (Normal o Tuberculosis) junto con su probabilidad en formato JSON. Esta arquitectura cliente-servidor aprovecha los recursos del backend y facilita el mantenimiento y la actualización centralizada, conforme a las mejores prácticas para la integración de modelos de deep learning en aplicaciones móviles.

El módulo de NLP funciona como un servicio independiente accesible mediante otra API REST. Este recibe las consultas en lenguaje natural y mediante RAG genera recomendaciones clínicas contextualizadas. El backend convierte la pregunta en un embedding semántico con SentenceTransformers, recupera fragmentos relevantes del corpus indexado en FAISS y genera la respuesta con el modelo DeepSeek R1. El resultado se envía en formato JSON para su integración y visualización en la interfaz móvil.

Ambos servicios, tanto para el diagnóstico médico como el de interpretación lingüística para recomendación, están desacoplados de la interfaz móvil como un microservicio, lo que proporciona flexibilidad, escalabilidad y la posibilidad de sustituir modelos individuales sin afectar al resto del sistema, siguiendo principios de diseño orientado a servicios y despliegues distribuidos para aplicaciones de con módulos de inteligencia artificial [41].

## IV. RESULTADOS

### A. Comparación de Modelos de CNN y Selección del Modelo más Eficiente para Diagnóstico

Los tres modelos CNN (MobileNetV2, ResNet50 y DenseNet121) fueron entrenados bajo condiciones homogéneas durante 40 épocas, utilizando el mismo conjunto de datos, parámetros y técnicas de optimización. El proceso de entrenamiento y validación requirió en promedio 3.2 horas por cada uno de los modelos. Los resultados obtenidos se resumen en la Tabla V, donde se presentan los valores de las métricas de desempeño para cada arquitectura y en la Fig. 6, que muestra su proceso de entrenamiento y validación.

MobileNetV2 registró los valores de, un 98.2% en *Accuracy*, un 98.9% en *Precision*, *Recall* y *F1-Score*, y un 99.7% en *AUC-ROC*; sin embargo, como muestra la Fig. 6. e) pérdida y f) precisión, el modelo evidenció un sobreajuste a partir de la décima época, lo cual comprometió su capacidad de generalización y anticipo un rendimiento inestable en escenarios reales. Por este motivo, fue descartado para la etapa de implementación.

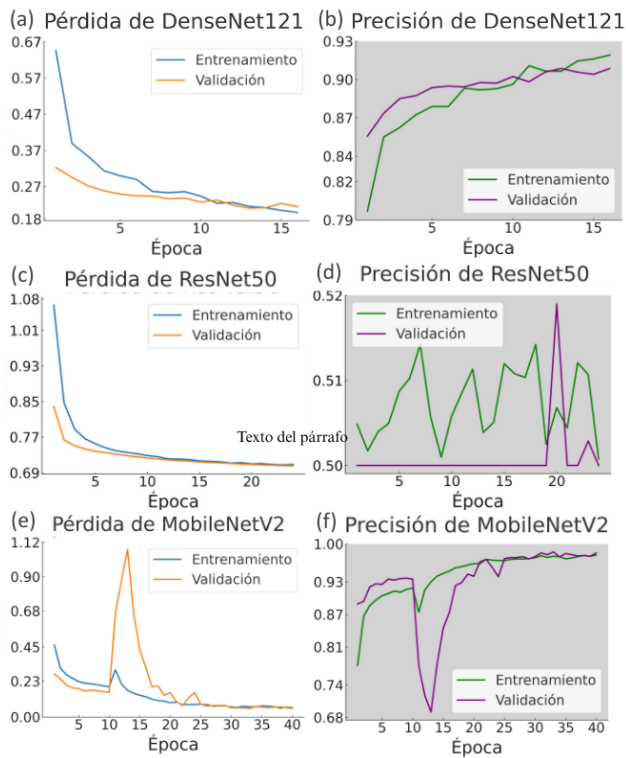


Fig. 6. Curvas de precisión y pérdida por época para los modelos evaluados.

ResNet50 presentó el rendimiento más bajo en todos modelos evaluados (Fig. 6.d), mostrando la menor capacidad de aprendizaje y estabilidad, lo que limitó su desempeño general frente las otras arquitecturas analizadas.

TABLA V

RESULTADOS COMPARATIVOS DE MÉTRICAS DE EVALUACIÓN PARA LOS MODELOS CNN

Métricas	MobileNetV2	DenseNet121	ResNet50
Accuracy	0.9816	0.9621	0.8483
Precision	0.9893	0.9685	0.7971
Recall	0.9893	0.9552	0.9345
Specificity	0.9862	0.9690	0.7621
Youden	0.9755	0.9242	0.6966
F1-Score	0.9893	0.9618	0.8603
ROC-AUC	0.9971	0.9943	0.9476

El modelo **DenseNet121** se consolidó como el modelo más equilibrado, alcanzando un *Accuracy* de 96.2%, *Precision* de 96.9%, *F1-score* de 96.2%, *Recall* de 95.5% y AUC-ROC del 99.4%. En el proceso de análisis, se priorizó la métrica de *Recall* sobre la precisión, con el propósito de minimizar los falsos negativos y evitar omitir casos reales de tuberculosis, logrando la mejor relación entre sensibilidad diagnóstica, robustez y estabilidad frente a las demás arquitecturas. Además, el análisis de las curvas de pérdida y precisión durante el proceso de entrenamiento y validación (Fig. 6 a y b, respectivamente) confirmó la consistencia y capacidad de generalización del modelo.

En contraste, del total de 290 imágenes correspondientes al conjunto de pruebas (clase Positivo), se detectaron 12 instancias catalogadas erróneamente como falsos negativos. El

análisis revela como causa de error al enmascaramiento por superposición anatómica y la insuficiencia de contraste en las radiografías evaluadas. Asimismo, la existencia de elementos de alta luminosidad altera la normalización y desvía la atención del modelo, fallando cuando se presentan geometrías atípicas que causan ambigüedad anatómica como se observa en la Fig. 7.

Por otro lado, del conjunto de pruebas de la clase Negativo, se encontraron con 11 instancias catalogadas como falsos positivos (Fig. 8). La clasificación presenta principalmente errores relacionados a elementos externos y de posicionamiento del paciente. Estas situaciones generan densidad adicional y bordes que el algoritmo interpreta de manera incorrecta como afectaciones pulmonares.

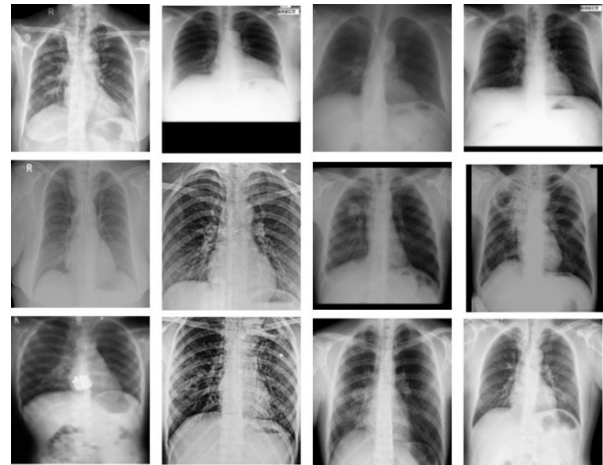


Fig. 7. Ejemplos de radiografías de tórax correspondientes a falsos negativos en la detección de tuberculosis.

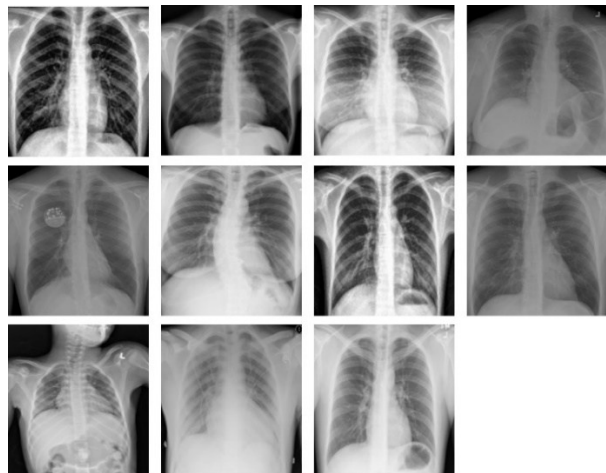


Fig. 8. Ejemplos de radiografías de tórax correspondientes a falsos positivos en la detección de tuberculosis.

### B. Resultados del Desarrollo del Asistente de Procesamiento de Lenguaje Natural

El modelo *DeepSeek-R1 free model variant* fue evaluado a través de la plataforma OpenRouter, enviando cada consulta en distintas cantidades de *chunks* (fragmentos de texto) y midiendo tanto el tiempo total de procesamiento en el *backend* como el objetivo de medir el tiempo de respuesta del modelo (Fig. 9). Los resultados obtenidos evidenciaron que el tiempo

promedio de respuesta aumenta proporcionalmente al número de *chunks* utilizados, especialmente en aquellas consultas que contienen un mayor volumen de información o número elevado de tokens generados.

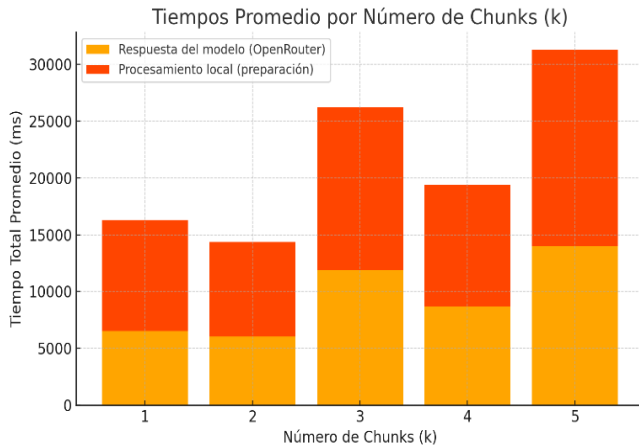


Fig. 9. Tiempo promedio de respuesta por el número de Chunks ingresados por el asistente conversacional.

En la comparación entre *backend* y OpenRouter, se observó que el primero añade un retardo adicional de 2 a 6 segundos por consulta debido al procesamiento propio y la gestión asincrónica. Además, el uso de múltiples *chunks* por pregunta incrementa el número total de solicitudes; esto tiene implicaciones directas en los límites de uso establecidos por OpenRouter y el modelo DeepSeek-R1 (20 solicitudes por minuto y, entre 50 y 1,000 diarias, según créditos disponibles).

### C. Resultados de la Interfaz de Usuario de la Aplicación Móvil

Se desarrolló una interfaz móvil funcional e intuitiva que facilita la interacción entre el usuario y los módulos de diagnóstico y recomendación. En el módulo de entrada (Fig. 10.a), el usuario puede cargar o reemplazar una radiografía de tórax; el sistema presenta advertencias para garantizar que la imagen sea clara y muestre la región pulmonar completa. Además, se muestran los formatos de archivo admitidos y se realizan validaciones internas para confirmar que sea una radiografía de tórax antes de proceder con el diagnóstico.

Una vez cargada la imagen, el procesamiento se inicia mediante el botón de Continuar. En el **módulo de respuesta** (Fig. 10.b), presenta la probabilidad estimada para tuberculosis mediante un indicador gráfico, acompañado de una recomendación inicial basado en el porcentaje obtenido. Adicionalmente, ofrece al usuario la posibilidad de consultar información complementaria sobre tratamientos, cuidados o aspectos generales del caso. volumen de información o número elevado de tokens generados.

Por su parte, el **módulo de asistencia diagnóstica** (Fig. 11) integra técnicas de NLP con el modelo DeepSeek R1 como núcleo generador de respuestas. El módulo emplea el resultado diagnóstico como prompt inicial para proporcionar explicaciones, apoyándose en una base de datos previamente indexada como base de conocimiento que permite interpretar consultas y generar respuestas coherentes, contextualizadas y adaptadas al ámbito clínico, priorizando que los usuarios

consulten con un especialista en todo momento y que la herramienta no reemplaza a una consulta médica profesional.

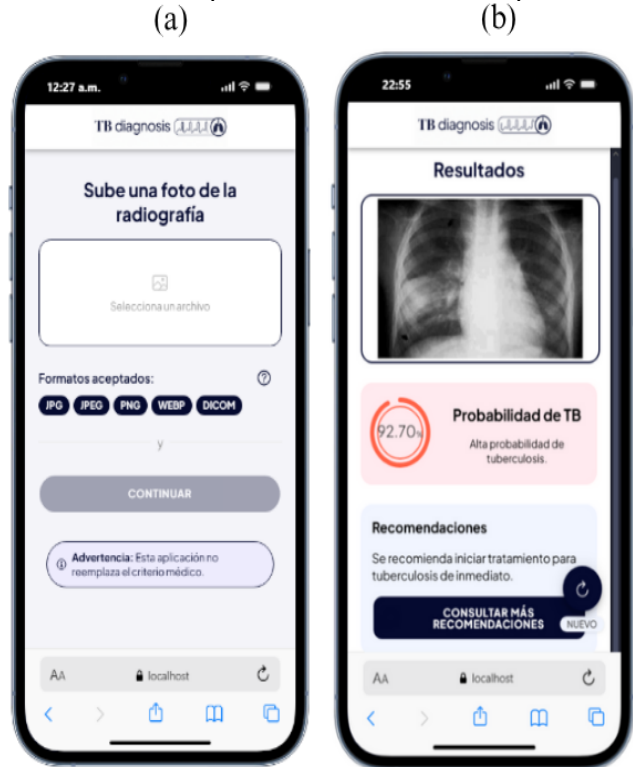


Fig. 10. Interfaz del prototipo de aplicación móvil desarrollada para el análisis y diagnóstico de tuberculosis.

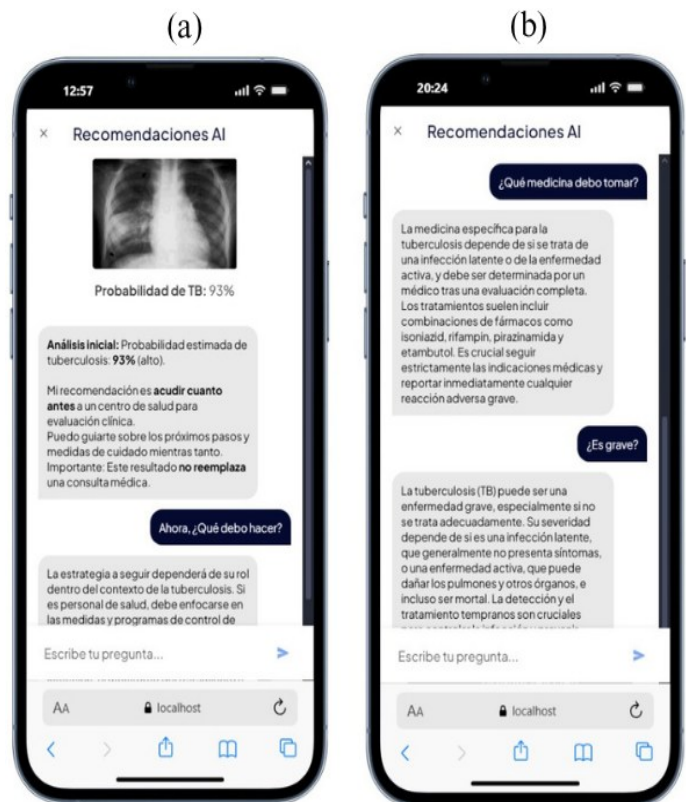


Fig. 11. Asistente conversacional integrado en la aplicación móvil para apoyar el proceso de detección de tuberculosis.

D. Resultados de Evaluación del Prototipo de Aplicación Móvil en un Entorno de Trabajo Real

Se aplicó un cuestionario (Tabla VI) a 12 médicos generales y especialistas en enfermedades neumológicas infecciosas, con el propósito de evaluar la utilidad, precisión diagnóstica y adecuación lingüística del sistema, en diversos centros de salud de la ciudad de Machala, Ecuador.

TABLA VI

PREGUNTAS REALIZADAS EN EL PROCESO DE VALIDACIÓN

ID	Preguntas de Cuestionario para Médicos
P1	¿La aplicación determina la enfermedad de tuberculosis con un nivel de probabilidad confiable?
P2	¿La aplicación es una herramienta útil de apoyo para los especialistas en el diagnóstico de tuberculosis y, proporciona información clara y comprensible sobre la enfermedad?
P3	¿La aplicación es usable, intuitiva y de fácil interpretación de los resultados?
P4	¿El chatbot proporciona recomendaciones clínicas claras, pertinentes y útiles ante un posible diagnóstico positivo de tuberculosis?
P5	¿El lenguaje usado por el chatbot es clínicamente adecuado y empático?
P6	¿Recomendaría esta aplicación como herramienta de apoyo en el diagnóstico de tuberculosis?

Los resultados revelan una percepción mayoritariamente “aceptable” de la aplicación, considerando valoraciones “De acuerdo” y “Totalmente de acuerdo” (Fig. 12). La confiabilidad diagnóstica (P1) obtuvo un 75% de aceptación, mientras que la utilidad como herramienta de apoyo y claridad informativa (P2) alcanzó un 66.7%. En la usabilidad y facilidad de interpretación (P3), el 83.4% consideró adecuada. En las recomendaciones clínicas del chatbot (P4) y en el lenguaje clínico y empatía del sistema (P5) obtuvo un 75% de aceptación.

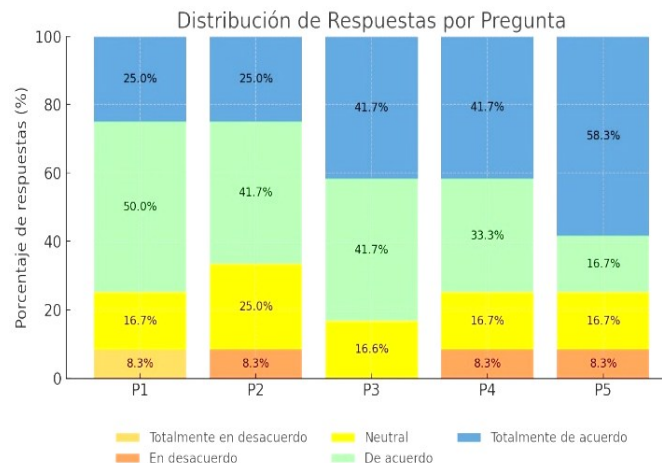


Fig. 12. Resultados de valoración del cuestionario de preguntas P1 a P5.

Finalmente, si recomendaría esta aplicación como herramienta de apoyo en el diagnóstico de tuberculosis (Fig. 13, P6), el 91.7% afirmó que sí. Esto evidencia pertinencia de las sugerencias generadas. Los resultados obtenidos respaldan la utilidad y viabilidad de la herramienta en contextos clínicos, demostrando su potencial como herramienta de apoyo

diagnóstico con oportunidad de optimización y mejora en la interacción lingüística.



Fig. 13. Resultados P6. ¿Recomendaría esta aplicación como herramienta de apoyo en el diagnóstico de tuberculosis?

VI. COMPARACIÓN CON LOS RESULTADOS DE OTROS ESTUDIOS

El uso de redes neuronales convolucionales para la identificación de enfermedades ha impulsado un creciente interés en su aplicación en múltiples entornos médicos gracias a su capacidad de apreciar patrones complejos que suelen ser omitidos por el factor humano, tomando fuerza como herramientas de apoyo en el área de salud.

En la siguiente sección se presenta un análisis comparativo entre los resultados obtenidos en este estudio y aquellos reportados en investigaciones previas sobre la detección automática de tuberculosis a partir de imágenes radiográficas de tórax. La Tabla VII reúne las principales características de los enfoques analizados, las arquitecturas implementadas en cada estudio y el mejor desempeño reportado en cada uno. Esta comparación permite reconocer los puntos en común entre los modelos CNN empleados en para su entrenamiento y validación, para establecer el posicionamiento del presente trabajo.

TABLA VII

TABLA COMPARATIVA DE ESTUDIOS BASADOS EN CNN PARA LA DETECCIÓN DE TUBERCULOSIS EN RADIOGRAFÍAS DE TÓRAX

Autores	#Imágenes	Modelos utilizados	Mejore desempeño
Rahman et al. [5]	7,000	ResNet101 DenseNet201 CheXNet InceptionV3 VGG19	DenseNet201: Accuracy 98.6%, Precision 98.57%, Recall 98.56%
Kotei & Thirunavukarasu [6]	3,600	DeiT- T+ResNet16	Accuracy 99.38%, Precision 99.24%, Recall. 99.49%
Guanca et al. [9]	564	Random Forest SVM CNN-Simple	Random Forest: Recall 77%, Specificity 80%
Trabajo propuesto	11,600	MobileNetV2 DenseNet121 ResNet50	<b>DenseNet121:</b> Accuracy 96.2%, Precision 96.85%, Recall 95.5%

VII. CONCLUSIONES

El estudio comparativo de arquitecturas CNN determinó que DenseNet121 constituye la arquitectura más equilibrada para la detección de tuberculosis, alcanzando un rendimiento del 96.2% en Accuracy, 95.5% en Recall y 96.2% en F1-score. Si bien el modelo demostró una alta robustez general, el análisis de errores entre sus falsos positivos y negativos, reveló unas

limitaciones circunstanciales en escenarios de alta complejidad caracterizados por superposición, insuficiencia de contraste o presencia de artefactos externos. Estos hallazgos validan la selección de la arquitectura DenseNet121 no solo por su precisión, sino por su estabilidad frente a factores variables en las radiografías analizadas.

Por otro lado, como aporte a la brecha tecnológica, se desarrolló una aplicación móvil bajo la metodología XP e implementada bajo una arquitectura cliente-servidor mediante API RESTful, incorporando una interfaz intuitiva y módulos diferenciados para las funciones de diagnóstico automatizado, la emisión de recomendación clínica y la interacción conversacional. El componente NLP, basado en el modelo DeepSeek-R1 mediante la plataforma OpenRouter, permitió generar respuestas clínicas correctas bajo el contexto del usuario.

Finalmente, la validación con profesionales de la salud reflejó niveles de aceptación positivos en términos de confiabilidad diagnóstica y calidad de interacción. Los niveles de aprobación del 75% al 83.4% relacionados a los aspectos de confiabilidad diagnóstica, usabilidad, facilidad de interpretación, la pertinencia y la claridad de las recomendaciones clínicas con empatía en su tono de respuesta por el sistema conversacional. Asimismo, el 91.7% de profesionales de la salud manifestó que recomendaría la aplicación como herramienta de apoyo en el diagnóstico de tuberculosis. Estos resultados confirman su viabilidad clínica y su potencial en contextos de recursos limitados.

#### REFERENCIAS

- [1] F. Estaji, A. Kamali, and M. Keikha, "Strengthening the global Response to Tuberculosis: Insights from the 2024 WHO global TB report," *J Clin Tuberc Other Mycobact Dis*, vol. 39, p. 100522, May 2025, doi: 10.1016/J.JCTUBE.2025.100522.
- [2] D. C. Falcón Córdova, J. C. Carrasquel Herrera, K. F. Viteri Tigse, H. P. Velasco Molina, and L. D. Sánchez Jaya, "Tuberculosis en el mundo y en el Ecuador, en la actualidad (2021)," *LATAM Rev. Latinoam. Cienc. Soc. Humanid.*, vol. 4, no. 6, pp. 658–673, Dec. 2023, doi: 10.56712/latam.v4i6.1476.
- [3] M. Z. Apolo Bravo and D. P. Andrade Campoverdeo, "Mycobacterium tuberculosis resistente a Rifampicina en la provincia de El Oro-Ecuador," *Revista Vive*, vol. 7, no. 19, Jan. 2024, doi: 10.33996/revistavive.v7i19.283.
- [4] L. S. Sánchez Figueroa, V. A. Guillén Muñoz, J. D. Pérez Pérez, P. A. Rodríguez Abrego, C. M. Caprile Mata, and K. L. Cartagena López, "Actualización en la detección de la tuberculosis a través de pruebas moleculares," *Alerta, Revista científica del Instituto Nacional de Salud*, vol. 7, no. 2, pp. 184–190, Jul. 2024, doi: 10.5377/alerta.v7i2.17129.
- [5] T. Rahman *et al.*, "Reliable Tuberculosis Detection Using Chest X-Ray With Deep Learning, Segmentation and Visualization," *IEEE Access*, vol. 8, pp. 191586–191601, Oct. 2020, doi: 10.1109/access.2020.3031384.
- [6] E. Kotei and R. Thirunavukarasu, "Tuberculosis Detection From Chest X-Ray Image Modalities Based on Transformer and Convolutional Neural Network," *IEEE Access*, vol. 12, pp. 97417–97427, Jul. 2024, doi: 10.1109/ACCESS.2024.3428446.
- [7] E. Kotei and R. Thirunavukarasu, "A Comprehensive Review on Advancement in Deep Learning Techniques for Automatic Detection of Tuberculosis from Chest X-ray Images," *Archives of Computational Methods in Engineering* 2023 31:1, vol. 31, no. 1, pp. 455–474, Aug. 2023, doi: 10.1007/S11831-023-09987-W.
- [8] G. Ramírez-Alonso, O. Prieto-Ordaz, R. López-Santillan, and M. Montes-Y-Gómez, "Medical Report Generation through Radiology Images: An Overview," *IEEE Latin America Transactions*, vol. 20, no. 6, pp. 986–999, Jun. 2022, doi: 10.1109/TLA.2022.9757742.
- [9] D. A. Guanca, C. J. Izetta, N. M. Pérez, and A. M. Verazay, "Diagnóstico Automático de Tuberculosis en Radiografías de Tórax aplicando Aprendizaje Automático," *Investigaciones en Facultades de Ingeniería del NOA*, pp. 108–119, 2023.
- [10] G. Y. Angulo Pupo, L. I. Saraiba Nuñez, and M. R. Moreno Pino, "Toma de decisiones con inteligencia artificial en el diagnóstico del paciente neonatal," *RILCO DS: Revista de Desarrollo sustentable, Negocios, Emprendimiento y Educación*, vol. 6, no. 62, pp. 80–88, Dec. 2024, doi: 10.51896/rilcods.v6i62.734.
- [11] A. K. Guerrero-Solís, G. F. Ruiz-Muñoz, D. A. Yépez-González, and M. N. Sánchez-Lascano, "El impacto de la inteligencia artificial en la producción científica," *Multidiscip. Lat. Am. J.*, vol. 3, no. 1, pp. 629–649, Apr. 2025, doi: 10.62131/mlaj-v3-n1-031.
- [12] N. Jiménez, S. Orellana, B. Mazon-Olivo, W. Rivas-Asanza, and I. Ramírez-Morales, "Detection of Leaf Diseases in Banana Crops Using Deep Learning Techniques," *AI 2025, Vol. 6, Page 61*, vol. 6, no. 3, p. 61, Mar. 2025, doi: 10.3390/ai6030061.
- [13] B. S. Guncay Saldaña, B. S. Guevara Pando, B. E. Mazón Olivo, and W. B. Rivas Asanza, "Diagnóstico de enfermedades bucales en imágenes basado en machine learning," *Revista Científica Arbitrada Multidisciplinaria PENTACIENCIAS*, vol. 7, no. 3, pp. 176–195, Mar. 2025, doi: 10.59169/pentaciencias.v7i3.1484.
- [14] H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, "Transfer learning for medical image classification: a literature review," *BMC Medical Imaging* 2022 22:1, vol. 22, no. 1, pp. 1–13, Apr. 2022, doi: 10.1186/S12880-022-00793-7.
- [15] A. W. Salehi *et al.*, "A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope," *Sustainability* 2023, Vol. 15, Page 5930, vol. 15, no. 7, p. 5930, Mar. 2023, doi: 10.3390/SU15075930.
- [16] S. Jacobo-Zavaleta and J. Zavaleta, "A Deep Learning Approach for Epilepsy Seizure Identification Using Electroencephalogram Signals: A Preliminary Study," *IEEE Latin America Transactions*, vol. 21, no. 3, pp. 419–426, Mar. 2023, doi: 10.1109/TLA.2023.10068845.
- [17] G. K. Sun and A. P. Ambrosy, "Applying Natural Language Processing to Electronic Health Record Data—From Text to Triage," *JAMA Netw Open*, vol. 7, no. 11, pp. e2443934–e2443934, Nov. 2024, doi: 10.1001/JAMANETWORKOPEN.2024.43934.
- [18] J. A. Calderón Blas, M. A. Cerdán, A. J. Sánchez García, and S. Domínguez Isidro, "Inteligencia Artificial en Sistemas de recomendación Médica: Una revisión de la Literatura," *ReCIBE, Revista electrónica de Computación, Informática, Biomédica y Electrónica*, vol. 12, no. 2, pp. C3–20, Nov. 2023, doi: 10.32870/recibe.v12i2.311.
- [19] D. F. Terraza Arciniegas, M. Amaya, A. Piedrahita Carvajal, P. A. Rodríguez-Marin, L. Duque-Munoz, and J. D. Martínez-Vargas, "Students' Attention Monitoring System in Learning Environments based on Artificial Intelligence," *IEEE Latin America Transactions*, vol. 20, no. 1, pp. 126–132, Jan. 2022, doi: 10.1109/TLA.2022.9662181.

- [20] Z. Jiang *et al.*, “Active Retrieval Augmented Generation,” *EMNLP 2023 - 2023 Conference on Empirical Methods in Natural Language Processing, Proceedings*, pp. 7969–7992, 2023, doi: 10.18653/V1/2023.EMNLP-MAIN.495.
- [21] C. Schröer, F. Kruse, and J. M. Gómez, “A systematic literature review on applying CRISP-DM process model,” *Procedia Comput Sci*, vol. 181, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.
- [22] J. Tamayo Abanto and B. A. Sauñe Huaman, “Soluciones de TI basadas en machine learning aplicables a problemas del sistema de salud peruano,” *BIOTECH AND ENGINEERING*, vol. 4, no. 1, pp. 156–163, Feb. 2024, doi: 10.52248/eb.Vol4Iss1.116.
- [23] S. G. Tetteh, “Empirical Study of Agile Software Development Methodologies: A Comparative Analysis,” *Asian Journal of Research in Computer Science*, vol. 17, no. 5, pp. 30–42, Feb. 2024, doi: 10.9734/ajrcos/2024/v17i5436.
- [24] S. Kiran and I. Jabeen, “Tuberculosis Chest X-rays Images,” Kaggle. Accessed: May 10, 2025. [Online]. Available: <https://www.kaggle.com/dsv/10304386>
- [25] Madyono and A. Nabilah, “Detection of Tuberculosis Disease in Lung X-ray Images Using the DenseNet121 Method,” *International Journal of Engineering, Science and Information Technology*, vol. 5, no. 2, pp. 379–385, Mar. 2025, doi: 10.52088/IJESTY.V5I2.853.
- [26] E. Goceri, “Medical image data augmentation: techniques, comparisons and interpretations,” *Artif Intell Rev*, vol. 56, no. 11, p. 1, Nov. 2023, doi: 10.1007/S10462-023-10453-Z.
- [27] O. Rainio, J. Teuho, and R. Klén, “Evaluation metrics and statistical tests for machine learning,” *Sci Rep*, vol. 14, no. 1, p. 6086, Dec. 2024, doi: 10.1038/s41598-024-56706-x.
- [28] G. M. Foody, “Challenges in the real world use of classification accuracy metrics: From recall and precision to the Matthews correlation coefficient,” *PLoS One*, vol. 18, no. 10, p. e0291908, Oct. 2023, doi: 10.1371/journal.pone.0291908.
- [29] J. Kivimäki, J. Bialek, W. Kuberski, and J. K. Nurminen, “Performance Estimation in Binary Classification Using Calibrated Confidence,” May 2025, doi: 10.48550/arXiv.2505.05295.
- [30] M. Madyono and A. Nabilah, “Detection of Tuberculosis Disease in Lung X-ray Images Using the DenseNet121 Method,” *International Journal of Engineering, Science and Information Technology*, vol. 5, no. 2, pp. 379–385, Mar. 2025, doi: 10.52088/ijesty.v5i2.853.
- [31] M. A. Hadi and H. I. Ali, “Control of COVID-19 system using a novel nonlinear robust control algorithm,” *Biomed Signal Process Control*, vol. 64, p. 102317, Feb. 2021, doi: 10.1016/j.bspc.2020.102317.
- [32] A.-L. Brun, M.-L. Chabi, C. Picard, F. Mellot, and P. A. Grenier, “Lung Transplantation: CT Assessment of Chronic Lung Allograft Dysfunction (CLAD),” *Diagnostics*, vol. 11, no. 5, p. 817, Apr. 2021, doi: 10.3390/diagnostics11050817.
- [33] World Health Organization, *WHO consolidated guidelines on tuberculosis. Module 1: prevention – tuberculosis preventive treatment*, Second. Geneva, 2020.
- [34] Centers for Disease Control and Prevention, *Core Curriculum on Tuberculosis: What the Clinician Should Know*, Seventh. 2021.
- [35] M. Kamil and D. Cakir, “Advances in Transformer-Based Semantic Search: Techniques, Benchmarks, and Future Directions,” *Turk. J. Math. Comput. Sci.*, vol. 17, no. 1, pp. 145–166, 2025, doi: 10.47000/tjmcs.1633092.
- [36] S. Y. Luis, D. G. Reina, and S. T. Marín, “Towards a Retrieval-Augmented Generation Framework for Originality Evaluation in Projects-Based Learning Classrooms,” *Education Sciences 2025, Vol. 15, Page 706*, vol. 15, no. 6, p. 706, Jun. 2025, doi: 10.3390/educsci15060706.
- [37] B. Liu *et al.*, “Improving Medical Vision-Language Contrastive Pretraining With Semantics-Aware Triage,” *IEEE Trans Med Imaging*, vol. 42, no. 12, pp. 3579–3589, Dec. 2023, doi: 10.1109/TMI.2023.3294980.
- [38] O. Baniş, D. Florea, R. Gyalai, and D. I. Curiaş, “Automated Specification-Based Testing of REST APIs,” *Sensors 2021, Vol. 21, Page 5375*, vol. 21, no. 16, p. 5375, Aug. 2021, doi: 10.3390/S21165375.
- [39] M. Rodríguez-Contreras, J. P. Sánchez-Solís, G. Rivera, and R. Florencia, “Framework to Support Radiologist Personnel in the Diagnosis of Diseases in Medical Images Using Deep Learning and Personalized DICOM Tags,” *Computación y Sistemas*, vol. 28, no. 3, pp. 1321–1348, Sep. 2024, doi: 10.13053/cys-28-3-5186.
- [40] T. Diot-Dejonghe *et al.*, “Development of a Secure Web-Based Medical Imaging Analysis Platform: The AWESOMME Project,” *Journal of imaging informatics in medicine*, vol. 37, no. 5, pp. 2612–2626, Oct. 2024, doi: 10.1007/S10278-024-01110-0.
- [41] C. Lee, H. F. Kim, and B. G. Lee, “A Systematic Literature Review on the Strategic Shift to Cloud ERP: Leveraging Microservice Architecture and MSPs for Resilience and Agility,” *Electronics 2024, Vol. 13, Page 2885*, vol. 13, no. 14, p. 2885, Jul. 2024, doi: 10.3390/ELECTRONICS13142885.



**Camily Bravo-Flores** holds a degree in Information Technology Engineering from Universidad Técnica de Machala. She has collaborated in the development of several IT solutions. Her interests include software engineering, user experience, artificial intelligence, deep learning, and natural language processing.



**Derik Aranda-Neira** holds a degree in Information Technology Engineering from Universidad Técnica de Machala. He has participated in the development of several IT projects, specializing in software engineering, user interface design, and database management, and has integrated various artificial intelligence technologies.



**Wilmer Rivas-Asanza** holds a degree in Computer Engineering from Universidad Católica de Cuenca. He has a diploma in IT Auditing and a diploma in Artificial Intelligence. He holds a master's degree in Teaching and Management in Higher Education from Universidad Estatal de Guayaquil, and a master's degree in Strategic IT Management from

Universidad Estatal de Cuenca. He earned a PhD in Information and Communication Technologies from Universidad de Coruña, Spain. He has professional experience in both the public and private sectors. He is a tenured professor at Universidad Técnica de Machala. His research interests include the strategic management of information technologies and artificial intelligence. He currently has several publications and indexed books that reflect his academic and professional career.



**Bertha Mazon-Olivo** holds a PhD in Information and Communication Technologies from Universidad de A Coruña, Spain. She is a research professor at Universidad Técnica de Machala, Ecuador and she earned a master's degree in Applied Informatics and a Bachelor's degree in Systems Engineering from Escuela Superior

Politécnica de Chimborazo. She also holds a master's degree in Big Data and Data Science from Universidad Internacional de Valencia, Spain. She is a member of the AutoMathTIC research group. She has participated in several research projects as director, co-director, and team member. Her research interests include Data Science, Big Data, Machine Learning, Deep Learning, and the Internet of Things. She has authored several publications, including indexed journal articles, books, book chapters, and conference papers presented at international events.